

研究とワープロ，パソコン，インターネット

大 串 夏 身

記憶装置はカセットテープ

研究の道具として，パソコン，インターネットは欠かせないものとなったが，私が研究を始めた頃は，ノート，カードの時代だった。

先行研究論文は，国立国会図書館の『雑誌記事索引』や学会誌掲載の論文の注，文献リストなどから必要なものをカードにとっておく，研究の素材も論文や資料集などから必要な部分をカードやノートにメモする，また，論文全体をコピーにとってバインダーにとじておくという時代だった。

1970年代末にノート型ワープロが売り出される。図書館の先輩が早速購入して職場に持ってきて，どのように使うのか皆の前で実演してくれた。記憶装置がカセットテープだった。文書を作って記憶装置に記録しようとするときに衝撃を与えてはいけない，衝撃を与える，例えば，机を強く叩いたりするとデータが記録されないことがある，という説明だった。長い文章をいっぺんに変換することはできなかった。しばらくすると，フロッピーディスクに記録するノート型ワープロが売り出されて，同室の上司が購入してきて，文書を作るのに使っていた。印刷機が付いていて，作った文書を印刷して見せてくれたときは，すごい機械だなと素直に感動した。

区名を半分しか変換できないワープロ

1980年4月に，それまで勤務していた東京都立日比谷図書館を出て，東京都23特別区が共同して作った財団法人特別区協議会に勤務することになった。2年目に，卓上に置く大型のワープロが1台持ち込まれた。業者が説明にきた。職員が会議室に集まり，業者が使い方を説明してくれた。文書を作りフロッピーディスクに保存して，それを印刷する。流れるように作業が進む，これは便利だ，と皆感心していた。ところが，業者に代わって職員が仕事の文

章を入力し始めると，なかなかうまくいかない。理由は，東京23特別区のうち12区しか区名がスムーズに変換できないのだ。あとの13区は1字，1字変換することになった。行政の専門用語はほとんどが1字ずつ変換する状態だった。さすがに皆でため息をついた。区名や地名，専門用語はその都度登録していけば変換効率は向上するという説明だった。実際にその機械を使うことになったが，ある時，置いてある部屋を覗いたら登録した専門用語が壁一面貼ってあった。皆で使うので，どういう言葉が登録されたか，あとの人が分かるようにということだった。面倒だなと思うとともに，このように変換辞書を作っていけば，効率良く文書を作れるようになるのだとも思った。

住所録で先行研究論文管理

最初にノート型ワープロを買ったのは1985年頃だと記憶している。ワープロに入っていた住所録のソフトを使って先行研究論文を管理することにした。住所録は，郵便番号，名前，住所，電話番号，備考などの項目があって，それぞれに一定の文字数を入力できる。名前の項目に著者名，住所の項目に論文タイトルと掲載雑誌名，電話番号に巻号（通号），ページ数，備考に刊年月を入力すると決めて，データを入力した。入力した文字列に対して検索ができる，著者名の並び替えもできる，並び替えたものを一覧表で出力もできる，という便利なものだった。

しかし，この試みはすぐに壁にあたった。

先行研究論文には，雑誌論文だけでなく，論文集に収録されている論文もある，図書の章や節も，同じように入力しておきたい。それに，論文の論文名，著者名，掲載雑誌名，巻号（通号），刊年月，ページ数だけでなく，内容のメモも入力したくなる。メモはカードの場合，余白や裏に書いておけばいいし，字数が多くなれば，2枚目に書いておけばいい。住所録のソフトは，項目も限られているし，

各項目に入力できる字数も限られている。また、研究が進んでいくと、先行研究論文だけでなく、新聞記事、写真・図版、地図、統計などの原資料の管理もワープロでしたくなる。量が増えてくるとカードより、コンピュータで管理したほうが管理しやすい。入力したデータを使って特定テーマの文献目録を作りたい、原稿に転用したいという欲求が生まれる。これらに応えるには、住所録ソフトでは無理がある。しかし、パソコンを買って自分でプログラムを作るという時間も実力もない。そこで、ワープロ付属の住所録ソフトで、自由に項目を作ってデータを入力することができるものがないか、探すことにした。パソコンの雑誌で調べてみると、NEC（日本電気）がそうしたものを販売しているらしいという記事が目についた。

文豪ミニ7HGが大活躍

幸い NEC は新宿に相談窓口を開設していて、一人1時間、無料で相談に応じていた。早速、申込んで休暇を取って行ってみると、申込み時に詳しく相談内容を言っていたので、該当のワープロ（「文豪ミニ7HG」という機種だった）を用意してくれていて、相談員が付属の住所録ソフトの使い方を説明してくれた。自由設計可能で、項目名も自由につけることができ、それぞれの項目に入力する文字数も指定できるという。

卓上型ワープロで少し値段は高かったが、購入した。

2枚のフロッピーディスクドライブが付いていて、1枚はソフト用のフロッピーを挿入し、1枚はデータの記録用のフロッピーを挿入する。両方とも3.5インチのフロッピーディスクだった。1枚のフロッピーディスクに記録できるデータ量は限られていたから、入力件数が増えるとフロッピーが2枚、3枚と必要になる。入力した文字列を対象に検索ができる。記録用のフロッピーディスクが2枚、3枚となると、ひとつのキーワードで検索をして集合を作るためには、2回、3回と繰り返さなくてはならない。これは、面倒なことだったが、時間をかければいいことなので、それほど苦にならなかった。

この文豪ミニ7HGを使って、『江戸・東京学研究文献案内』（大串夏身/江戸・東京資料研究会編、青弓社、1991年）に収録した「ことば（件名）から探せる研究文献目録」（111ページ～240ページ）、「東京関係写真・銅版画索引」

（243ページ～322ページ）、「『東京市史稿』市街篇に見る調査一覧 慶応3年～明治10年」（325ページ～350ページ）などを作成した。また、各種レファレンスブックの書誌情報の管理にも使った。

「ことば（件名）から探せる研究文献目録」は、江戸東京関係の研究論文を、内容を示す「ことば（件名）」から検索できるようにしたもので、「ことば（件名）」を見出し語にして50音順に並べた。見出し語は、約1100だった。見出し語のもとに並べた論文はすべて現物を見て、内容をチェックしたものだった。そのために、勤務先の東京都立中央図書館の所蔵資料を夜、書庫で調べて、ある程度調べ終わると、つぎに母校の早稲田大学図書館に行き、卒業生でも書庫に入って自由に閲覧、コピーができたので、仕事が終わってから行って、9時過ぎの閉館まで書庫で調べた。週2、3日は早稲田大学図書館の書庫で文献を漁っていた。早稲田がある程度終わると、次は国立国会図書館。国立国会図書館でないと所蔵していない雑誌を重点的に調べた。当時、国立国会図書館は、5時閉館以降、一般研究室が8時まで開いていて、そこは研究テーマと研究内容を書いた書類を提出すると利用させてくれた。夜なので、利用者が少ない。書庫出納もしてくれたので、カウンターに請求票を出すと、しばらくすると資料が出てきた。昼間よりずっと早かった。たしか、図書、雑誌あわせて3冊だったと記憶している。ただ、私のように片っ端から請求する利用者には、出納担当のアルバイトの学生も請求のたびに取りに行くのが面倒らしく、「今日はどれ位見ますか、最初に請求票を出してもらえば準備しておきますよ」と声をかけてくれたので、そうしてもらった。随分と早く調べることができた。ノートにメモをとって、休みの日などにワープロに入力するという生活だった。

「東京関係写真・銅版画索引」は、31の写真集、図版集に入っている東京関係の写真・図版を1枚、1枚、「ことば（件名）」で探せるようにしたもので、たぶん、『江戸・東京学研究文献案内』では一番良く使われたものではないかと思う。

『『東京市史稿』市街篇に見る調査一覧 慶応3年～明治10年』は、『東京市史稿』市街篇に掲載されている各種調査を調べることができるよう、調査時期順に、調査名、概要、掲載市街篇巻数、掲載ページを記載したもので、取

り敢えず明治 10 年までの分を掲載した。東京ではさまざまな調査が行われていて、それらが研究の対象になっていないという問題がある。調査は、都市研究の基礎となるもので、これを系統的に明らかにしたいという意欲の現れだった。

続いて『江戸・東京学雑誌論文総覧』（大串夏身/江戸・東京資料研究会編、青弓社、1994 年）を出版した。これは「ことば（件名）」から江戸・東京関係の論文を探すことができるもので、620 ページ、収録件数 4,000 件を越えたが、作業の途中で「文豪ミニ 7HG」から NEC のノートパソコン 98 シリーズ、PC-9801 NS/T に変更している。新しい書誌情報は商用オンラインデータベースで検索する時代が来ていたからだ。

商用オンラインデータベースへ挑戦

1990 年になると商用オンラインデータベースが普及してくる。国立国会図書館の所蔵データや日外アソシエーツの雑誌論文データベースなどが商用オンラインデータベースとして販売されるようになった。これらを検索しなくては研究論文の迅速で十分な把握ができない。パソコンを NEC 製にしたのは、文豪ミニ 7HG で作ったデータを継続して使うために NEC のノートパソコンの方がいいと思ったからだ。(注 1)

フロッピーディスクドライブにソフト用のフロッピーを挿入して使うタイプだったと記憶している。卓上型ワープロとの違いは、使えるソフトの数が増えたことと、インターネットが使える、商用オンラインデータベースが検索できるようになったことだ。

使えるソフトが増えたということで良かったことは、キーワード抽出ソフトなどワープロでは使えないソフトが使って研究に生かすことができたことだ。キーワード抽出ソフトはジャストシステムが発売したものを購入して使ったが、これは見出し語の分析、論文のなかでの用語の使い方の分析、また中島みゆきの歌詞の分析などに使った。全文を対象にしてキーワードを抽出して 50 音順に並べてくれた。宮沢賢治の作品の分析にも使ってみたが、使っている言葉の特徴を把握する上で参考になった。ただ、ノートパソコンでは処理能力が低く、長い時間がかかった記憶がある。長時間作業をさせていると非常に熱を発するので、夏

の夜にはうちわで扇いだりしながら処理が終わるのを待っていた。カード型データベースは、アスキー発売の The CARD を使った。マイクロソフト製のカード型データベースより圧倒的に使いやすかったが、残念ながら Windows 95 に対応する製品は売りだされなかった。ワープロソフトはジャストシステムの一太郎を使ったが、一太郎を何枚かのフロッピーにコピーして、研究分野毎に差し替えて、専門分野の用語変換の効率を上げることができた。私の場合、江戸・東京、図書館、近代史、歌謡曲などの論文、図書を書いていたので、これは大変助かった。

商用オンラインデータベースへの接続は時間がかかった。パソコンを購入して、説明書に付いていたネットワークプロバイダーの入会案内に必要事項を記入して送った。すぐに ID、パスワードが書かれた書類が届いた。オンラインデータベースに接続するためのモデムの紹介などがあったので、家電量販店に行ってモデムを買ってきた（もちろん NEC 製だった）。

パソコンにつないで設定をして、電話回線につないだが、データの交信ができない。説明書に従っていろいろやってみるがうまくいかない。翌日、昼休み、NEC の相談窓口 に電話して聞いてみる。これを 2 回繰り返して、やっとつながった。

この当時、商用オンラインデータベースの検索は、検索項目を半角英文字で指定して、その文字のあとに、検索用の文字を入力するという方式だった（コマンド方式と呼んでいた）。終了する時も、特定の文字（多くは「E」だったが）を入力して、「enter」キーを押すという方法だった。最初、接続して検索した時は、本当に終了したのか不安になった。検索用のマニュアルを見ながらの操作だった。

会員となった NEC の子会社のネットワークプロバイダー（今は「Biglobe」となっている）は、いろいろな会員向けの有料のサービスを用意していた。講談社や日経 BP 社発行の雑誌の論文を 1 論文だけでもデジタル化されたものをネットワークを通して購入できた。帝国データバンクの会社情報や日本能率協会マーケティングデータバンクのデータベース、国立国会図書館の所蔵情報データベースなども検索できた。日外アソシエーツは個別に契約して検索した。商用オンラインデータベースだけでなく、インターネットも検索できた。検索は半角英文字を入力して検索する方法

で、画像があるページを表示するのは非常に時間がかかった。

『風俗画報』などの索引の作成

コンピュータを使ってデータを入力すると、そのデータ全部を対象に検索が迅速にできるという利点が生まれる。各種の索引を作れば、研究はより容易になる。そこで、江戸東京関係の索引を作ることにした。

写真集の索引、統計書の索引、地図の索引の3つを企画した。写真集は、図書館で写真のタイトルが書いてあるページをコピーして、写真を見ながら件名等をメモして、それをコンピュータに入力した。統計書は、図書館で『東京府統計書』、『警視庁統計書』などの主要統計書の目次等をコピーして入力した。統計書の索引の一部は、『東京府・市二次統計書データベース——書誌情報レベルの作成』として刊行した。これにはCD-ROMが付いていて、CD-ROMに収録してあるカンマ形式のテキストデータを使用するコンピュータにダウンロードすればデータベースになる、というものだった。(注2)

地図は、1912年に出版された『東京市及接續郡部地籍地図』上・下巻を対象にした。これは東京15区と周辺5郡の土地について所有者だけでなく、実際に使っている会社、行政組織、個人などを地図上に記載したもので、索引を作ればそれらが簡単に検索できる。(注3) 地籍地図の索引は、すべての地図から読み取った文字列をコンピュータに入力するところまで終ったが、校正が終っていないのでまだ公表できないでいる。

また、『風俗画報』のデジタル化(CD-ROM化)にともなう索引の作成も行った。(注4) 最初に構想した方法は、すべてのページをデジタル化して、デジタル化したものから、文字データをテキストデータとして抽出して、それに対してキーワード抽出ソフトでキーワードを抽出して、その中から、必要なキーワードを索引として選ぶというものだった。それぞれにソフトがあるので、それを動員すればあまり時間をかけずにできるという予想だった。しかし、デジタル化したものからテキストデータとして抽出するところがうまくいかなかった。古い文字に抽出ソフトが対応できない。そこで、すべてのページにあたって、索引用にキーワードを抽出することにした。大学院生をアルバイト

として雇用してもらって、入力作業をしてもらった。キーワードとして図版の中の描かれている物や花なども採用することにした。(注5)

インターネットの充実と新たな研究上の課題

2005年頃からインターネット上で無料で閲覧できるデジタル化された学術関係の情報量が多くなってきた。これは、1995年、ブリュッセルで開かれたG7電子通信閣僚会議(後に情報関係閣僚会議とも呼ばれるようになった)で合意された電子政府、電子商取引などのプロジェクトが実を結びはじめたからだと言ってもいい。資料のデジタル化とインターネットへの公開に関連するのは、電子政府、電子図書館、電子博物館のプロジェクトだった。日本はフランスとともに電子図書館プロジェクトの幹事国になって、電子図書館化を進めた。国内では、国立国会図書館、国立大学附属図書館、国文学研究資料館などが中心となって、所蔵資料のデジタル化とインターネットへの公開に取り組んだ。

電子博物館のプロジェクトで、大英博物館、ルーブル美術館など主要な博物館、美術館の所蔵資料のデジタル化とインターネットへの公開が進められた。大英博物館は600万点の資料のデジタル化を行ったという。電子政府のプロジェクトによって各国の政府の文書や統計書、各種調査報告書などがインターネットで閲覧できるようになった。

学術雑誌や研究成果などは商用オンラインデータベースで1980年代から提供されていたが、2000年前後から学術機関所蔵資料、研究成果のデジタル化とインターネットへの公開を目的とした「機関リポジトリ」の取組が各国で進められ、最新の研究成果もインターネットで読めるようになりはじめた。

さらに最新刊のものも含めた図書の大規模なデジタル化を進めてきたGoogleが、2004年10月「Google Print」という名称で検索・閲覧サービスの本格運用を開始した。(これは後に「Google Book Search」さらに「Google Books」と名称を変更している。)これによってかなりの図書(注6)が、全文を対象とした検索と閲覧ができるようになった。

こうしたことに加えて、インターネットで世界の情報が検索できるようになり、研究に必要な情報も入手できるようになった。私が研究領域としている図書館では、従来検

索・閲覧することが難しかった図書、研究論文が検索できるようになったし、各国の図書館のさまざまな試みが分かるようになった。日本国内に限定した特定のサービスに関する研究を行う場合でも、諸外国の図書館の同一あるいは類似のサービスやその研究成果などを検索して事前に知ることができるようになったし、それらを事前の研究成果、検討の資料として、研究のプランに組み込むことが必要となった。研究の国際化が進みつつあると言っているだろう。これからの研究者は、従来の方法に加えて、インターネットを活用した研究テーマに関する学術情報を世界規模で収集、管理する方法を確立する必要がある。これは新しい研究上の課題とも言える。

注1 NECは、ワープロで作ったデータをパソコンでも使えるようにする変換ソフトを売っていた。それを購入して、文豪ミニ7HGで作ったデータをパソコンで使えるようにした。

注2 東京都江戸東京博物館都市歴史研究室編、2000年（東京都江戸東京博物館調査報告書第9集）。統計書そのものを復刻することも企画した。国勢調査などを中心とした『近代都市統計資料集成：東京市市勢調査篇』をゆまに書房から1998年、私が監修して出版した。東京市市勢調査篇は全31巻で、大正9年、昭和5年の国勢調査の東京市及び周辺5郡の詳細統計表（「原表」と呼んでいる）と、同様の調査が明治41年、大正13年にも行われているので、その詳細統計表を復刻した。明治41年の調査は国勢調査の前段に行われた詳細な調査で、大正9年の国勢調査と調査項目が同じものが多い。大正13年の調査は関東大震災のあとに特に行われた調査で、今回の東日本大震災でにわかに注目された調査である。このあと労働統計、物価統計、産業統計など続編として出す予定だったが、実現しなかった。

注3 東京市関係の区分地図などはほとんど索引が作られているが、一番詳しいこの地図はまだ作られていなかった。2012年現在も作られていないと思う。この地図は、上・下巻2冊に『東京市及接續郡部地籍臺帳』がついて3冊セットになって出版されたものである。これは柏書房から縮小されて復刻されているが、もとの図版の状態が悪いため、地図上の文字がはっきりしないところがある。それに原版は彩色されているが、復刻版は無色である。私が調べた限りでは早稲田大学図書館所蔵のものが一番美しい。

注4 ゆまに書房が製作・発売する企画で、樋田満文先生（当時・武蔵野女子大学教授、故人）が指導して、私と横山泰子氏（現・法政大学教授）が実務を担当した。

注5 1997年に出版されたCD-ROM版を検索してもらおうと分かるのだが、タイトル中のキーワードにないものもかなり検索できるようになっている。ただし、図版の中の「モノ」が当時そのように呼ばれていたかを特定できないものも少なくなく、キーワードを付けることができなかったものもかなりあった。

注6 影山幸一「デジタルアーカイブ羅針盤1 デジタルアーカイブの歴史」（infocom作成サイトの内の「デジタルアーカイブシステムコラム」のページ（<http://www.infocom.co.jp/das/column/column/index2.html>）2012年9月8日閲覧）によれば、2010年の時点で「4,860,000件ヒットする」とある。正確な件数は公開されていないが、現在も新しい図書が追加されている。

（おおぐし なつみ 大学院生活機構研究科人間教育学専攻）